

Leveraging Pre-trained Large Language Models to Construct and Utilize World Models for Model-based Task Planning

Year: 2023 | Citations: 245 | Authors: L. Guan, Karthik Valmeekam, S. Sreedharan

Abstract

There is a growing interest in applying pre-trained large language models (LLMs) to planning problems. However, methods that use LLMs directly as planners are currently impractical due to several factors, including limited correctness of plans, strong reliance on feedback from interactions with simulators or even the actual environment, and the inefficiency in utilizing human feedback. In this work, we introduce a novel alternative paradigm that constructs an explicit world (domain) model in planning domain definition language (PDDL) and then uses it to plan with sound domain-independent planners. To address the fact that LLMs may not generate a fully functional PDDL model initially, we employ LLMs as an interface between PDDL and sources of corrective feedback, such as PDDL validators and humans. For users who lack a background in PDDL, we show that LLMs can translate PDDL into natural language and effectively encode corrective feedback back to the underlying domain model. Our framework not only enjoys the correctness guarantee offered by the external planners but also reduces human involvement by allowing users to correct domain models at the beginning, rather than inspecting and correcting (through interactive prompting) every generated plan as in previous work. On two IPC domains and a Household domain that is more complicated than commonly used benchmarks such as ALFWorld, we demonstrate that GPT-4 can be leveraged to produce high-quality PDDL models for over 40 actions, and the corrected PDDL models are then used to successfully solve 48 challenging planning tasks. Resources, including the source code, are released at: <https://guansuns.github.io/pages/llm-dm>.